

# Linear and Nonlinear Frequency-Division Multiplexing

Mansoor I. Yousefi and Xianhe Yangzhang

**Abstract**—Two signal multiplexing strategies for multi-user optical fiber communication are considered: Wavelength-division multiplexing (WDM) and nonlinear frequency-division multiplexing (NFDM), based on the nonlinear Fourier transform (NFT). The inverse NFT is implemented using an approach that is dual of the forward NFT and does not require solving integral equations. In contrast to prior work where NFDM is applied to single-user channels or combined with WDM, in this paper users' signals are multiplexed in the nonlinear Fourier domain. Furthermore, each user now sends a sequence of symbols. NFDM orthogonalizes the nonlinear Schrödinger equation in a generalized frequency and a generalized time. Thus all degrees-of-freedom are (deterministically) independent. As an example, a data rate of 10.5 bits per complex degree-of-freedom is achieved in a 15-user NFDM with 60 GHz overall bandwidth at average input power  $\mathcal{P} = -0.33$  dBm. Data rate in a comparable 15-user WDM with the same bandwidth and power is 5.26 bits per complex degree-of-freedom. The ratio of the spectral efficiency of NFDM and WDM is 2.23, taking into account bandwidth and time duration of signals at channel input and output. Data rates are plotted as a function of the average input power, indicating that the NFDM rate increases monotonically with transmit power, in contrast to the WDM rate which characteristically vanishes as transmit power is increased more than an optimal value. Thus, as a result of the nonlinear multiplexing of users' signals, NFDM spectral efficiencies higher than the corresponding WDM spectral efficiencies are demonstrated at high input powers.

**Index Terms**—Optical fiber networks, nonlinear Fourier transform, nonlinear frequency-division multiplexing, wavelength-division multiplexing, multi-user communication.

## I. INTRODUCTION

**T**HIS paper concerns data transmission using the nonlinear Fourier transform (NFT) [Part I–III] [2]–[4].

It was recently pointed out that one factor contributing to data rate limitation in current optical fiber systems is the communication method [Part III, Section II]. Based on the NFT, nonlinear frequency-division multiplexing (NFDM) was introduced to address this limitation [Part I–III]. In this paper, we make progress on NFDM.

In Sections III and IV, we review the origin of data rate limitation in current approaches in optical fiber networks, introduce the nonlinear Fourier transform and NFDM. The presentation is shortened to essentials.

Prior work illustrates how NFDM can be applied with preliminary data rates. However, an NFDM spectral efficiency higher than the corresponding one in a comparable

wavelength-division multiplexing (WDM) has not yet been demonstrated. To demonstrate high spectral efficiencies, we consider the nonlinear Schrödinger (NLS) equation in the defocusing regime. In this case, the operator underlying channel is self-adjoint, which simplifies computations. The analyticity of the  $\log a(\lambda)$ , where  $a(\lambda)$  is a nonlinear Fourier coefficient, can be exploited to perform the inverse NFT without solving integral equations. In Section V, we simplify, test and apply existing numerical methods in the defocusing regime. The forward and inverse numerical algorithms are presented as dual of one another, *i.e.*, essentially the same algorithms (as in the Fourier transform).

Multi-user data rates were presented in [Part III]. In particular, in Section V. D, “Spectral Efficiencies Achievable by Modulating the Continuous Spectrum,” information was modulated in the continuous spectrum for each user. However, in preliminary simulations [Part III, Fig. 9(b)], users' signals were still multiplexed linearly, so as to be able to carry out computations. As a result, NFDM achieved spectral efficiencies approximately equal to WDM rates. It was articulated that, in order to improve existing data rates, NFDM should eventually be applied to networks — where a “capacity crunch” occurs [5, Section 6.7. 1], [Part III, Sections II.C, V. D. and VI. A]. In Sections VI and VII, we take the next step and multiplex users' signals in a nonlinear manner as required in NFDM.

Users' signals are orthogonal in NFDM. To have orthogonal degrees-of-freedom within each user, a *guard time* is used between multi-solitons. In the defocusing regime, NFT is a one-dimensional function, similar to the Fourier transform. This allows, as in the additive white Gaussian noise (AWGN) channel, to pack users' pulses optimally in a *generalized time*, according to Nyquist-Shannon sampling theorem. As a result, the channel is (deterministically) orthogonalized in a nonlinear spectral, as well as a temporal domain. Thus, in the defocusing regime, the spectral efficiency loss due to guard times is avoided.

In Section VII, we push numerical algorithms to high power transmission and examine the influence of noise and stochastic interference. We continue multi-user simulations in [Part III, Fig. 9 (b)], with nonlinear multiplexing, and demonstrate NFDM data rates higher than WDM rates subject to the same constraints. Importantly, the NFDM rate increases monotonically with transmit power, in contrast to the WDM rate which characteristically vanishes as transmit power tends to infinity.

We review literature on using NFT for communication in Section IX. Here, we also comment on multi-soliton communi-

cation, and explain the way that NFT could usefully be applied to data transmission.

The above contributions bring NFDM closer to practice, so that, in the defocusing regime, complexity and numerical accuracy remain as limiting factors. Similar results are anticipated in the focusing regime. Defocusing regime was considered because of the computational tractability and similarity of the NFT and Fourier transform.

## II. NOTATION

Real, complex and natural numbers are denoted by  $\mathbb{R}$ ,  $\mathbb{C}$  and  $\mathbb{N}$ , respectively. The complex Banach space of functions  $f : \mathbb{R} \mapsto \mathbb{C}$  with  $p$ -norm is represented by  $L^p(\mathbb{R})$ . The Hilbert space of  $T$ -periodic functions with the inner product  $\langle f, g \rangle = \frac{1}{T} \int_0^T f(t)g^*(t)dt$  is shown by  $L^2([0, T])$ . The Fourier transform operator is  $\mathcal{F}$ , with convention in [8, Eq. 1]. Probability distribution of a complex Gaussian random variable with mean  $\mu$  and variance  $\sigma^2$  is denoted by  $\mathcal{N}_{\mathbb{C}}(\mu, \sigma^2)$ . Independent identically distributed is abbreviated as i.i.d.

## III. FIBER-OPTIC NETWORKS

In this section, we present the NLS model in a network environment and summarize the origin of data rate limitation in current approaches.

### A. Network NLS Model

Signal propagation in single-mode single-polarization optical fiber with distributed amplification can be modeled by the stochastic NLS equation. In normalized form, the equation reads [Part I, Eq. 3]

$$j \frac{\partial q}{\partial z} = \frac{\partial^2 q}{\partial t^2} - 2s|q|^2q + n(t, z),$$

where  $j = \sqrt{-1}$ ,  $q(t, z)$  is the complex envelope of the signal as a function of time  $t$  and distance  $z$  along the fiber,  $n(t, z)$  is white Gaussian noise with power spectral density  $\sigma_0^2$ , and  $s = \pm 1$ , where  $s = 1$  (respectively,  $s = -1$ ) corresponds to the defocusing (respectively, focusing) regime. The signal and noise are bandlimited to  $B$  Hz for all  $0 \leq z \leq \mathcal{L}$ . We refer the reader to [Part I, Section I] for further details about the model.

In this paper, we consider a *network environment*. This refers to a communication network with the following assumptions [Part III, Section II. B. 3]: (1) There are multiple transmitter (TX) and receiver (RX) pairs; (2) There are add-drop multiplexers (ADMs) in the network. The signal of the user-of-interest can co-propagate with the signals of the other users in part of the link. The location and the number of ADMs are unknown; (3) Each TX and RX pair does not know the incoming and outgoing signals in the path that connects them; see Fig. 1.

To make use of available bandwidth, data can be modulated in distinct wavelengths. As a result of using more bandwidth, data rates, measured in bits per second, were rapidly increased with the introduction of the wavelength-division multiplexing in 1990s. However, the communication efficiency of WDM, measured in bits/s/Hz, is low. From a communications point

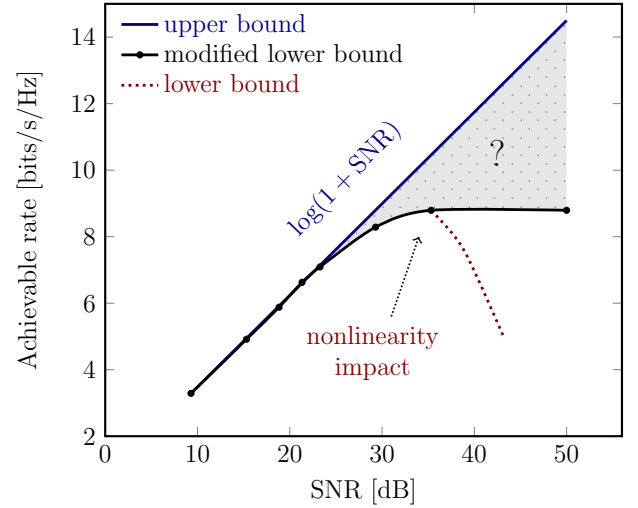


Fig. 2. Achievable rates of the WDM-BP strategy in a network environment.

of view, WDM underlies transmitter strategy. At the receiver, it is customary to apply some form of back-propagation (BP).

The achievable (per-user) rates of the WDM-BP strategy in a network environment are shown in Fig. 2. It can be seen that achievable rates (bits per complex degree-of-freedom) saturate (or decrease depending on assumptions), as input power is increased. Although numerous explanations exist in the literature, inter-channel interference and inter-symbol interference (ISI) resulting from nonlinearity [6] [Part III, Section II], as well strong filtering in the presence of spectral broadening in WDM [7, Section VIII], are the primary reasons for data rate limitation at high powers in WDM.

### B. Origin of Data Rate Limitation in Current Methods

It was understood in the past few years that nonlinear interactions do not limit capacity in deterministic integrable optical networks. These interactions arise from *methods of communication*, which disregard nonlinearity [Part I–III]. After abstracting away non-essential aspects, current methods, in essence, modulate linear-algebraic modes of transmission. These linear multiplexing methods include WDM, time-division multiplexing (TDM), orthogonal frequency-division multiplexing (OFDM), space-division multiplexing (SDM), and pulse-train transmission (see Fig. 3 (a)). When linear multiplexing is applied to nonlinear channels, interference and ISI are inevitable. In a network environment interference cannot be removed, while intra-channel ISI can *partially* be compensated for using signal processing. The idea of sharing bandwidth in WDM, and integration in SDM, conflict with nonlinearity. Since the latter is fixed by physics, it was proposed to replace the approach [Part I–III].

Yousefi and Kschischang recently proposed nonlinear frequency-division multiplexing which is fundamentally compatible with the channel [Part I–III]. NFDM is based on observation that the NLS equation supports nonlinear eigenfunctions which have a crucial independence property, the key to build a multi-user system. The tool necessary to reveal independent signal degrees-of-freedom is the NFT. Based

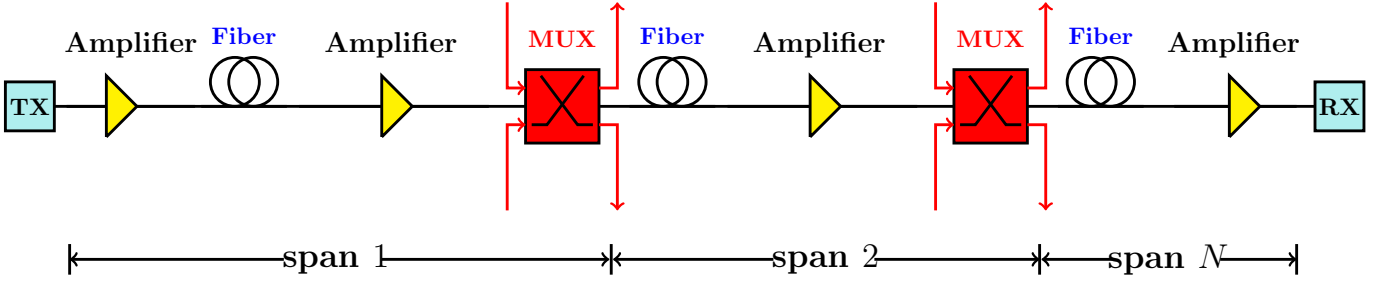


Fig. 1. Block diagram of a network environment. MUX denotes multiplexer, which in the current methods is ADM.

on that, NFDM was constructed which can be viewed as a generalization of OFDM to nonlinear optical fiber. Exploiting the integrability property, NFDM modulates non-interacting degrees-of-freedom. In the absence of noise, in contrast to WDM, achievable rate of NFDM is infinite (see Section VII).

#### IV. SUMMARY OF NFDM

In this section, we briefly review NFT and NFDM from [Part I–III].

Let  $T : \mathcal{H} \mapsto \mathcal{H}$  be a linear map on a separable complex Hilbert space  $\mathcal{H}$ . Consider the channel

$$y = T(x) + n, \quad (1)$$

where  $x$  and  $y$  are input and output signals and  $n$  is Gaussian noise on  $\mathcal{H}^1$ . The channel can be discretized by projecting signals and noise onto an orthonormal basis  $\{\phi_k\}_{k \in \mathbb{N}}$  of  $\mathcal{H}$

$$\{x, y, n\} = \sum_{k=1}^{\infty} \{x_k, y_k, n_k\} \phi_k, \quad (2)$$

where  $x_k, y_k, n_k \in \mathbb{C}$  are degrees-of-freedom. This results in a discrete model

$$y_k = x_k \langle T\phi_k, \phi_k \rangle + \underbrace{\sum_{i \neq k} x_i \langle T\phi_i, \phi_k \rangle}_{\text{linear interactions}} + n_k. \quad (3)$$

Depending on the basis, interactions in (3) could refer to inter-symbol interference in time, inter-channel interference in frequency, or generally interaction among degrees-of-freedom in any of the methods in Fig. 3(a).

Suppose that  $T$  is diagonalizable, *i.e.*, it has a set of eigenvectors forming an orthonormal basis of  $\mathcal{H}$  [2, Theorem 6]. In this basis, interactions in (3) are zero and

$$y_k = \lambda_k x_k + n_k, \quad (4)$$

where  $\lambda_k$  is an eigenvalue of  $T$ . As a result, the channel is decomposed into parallel independent scalar channels.

Interactions in (3) arise if the basis used for communication is not compatible with the channel. As a special case, let  $\mathcal{H} = L^2([0, L])$  and  $T$  be the convolution map  $T_h(x) = h(t) * x(t)$ , where  $h(t) \in L^1(\mathbb{R})$  is channel filter and  $*$  denotes convolution. The eigenvectors and eigenvalues of  $T_h$  are

$$\phi_k(t) = \frac{1}{\sqrt{L}} \exp(-jk\omega_0 t), \quad \omega_0 = \frac{2\pi}{L},$$

<sup>1</sup>Defined via projections (2), with  $n_k \sim \text{i.i.d. } \mathcal{N}_{\mathbb{C}}(0, 1)$ .

and  $\lambda_k = \mathcal{F}(h)(k\omega_0)$ . The Fourier transform maps convolution to a multiplication operator according to (4), where  $x_k, y_k$  and  $n_k$  are Fourier series coefficients. Interference and ISI are absent in the Fourier basis. In OFDM, information is encoded in independent spectral amplitudes  $x_k$ .

We explain NFDM in analogy with OFDM. First, we define NFT as follows. Consider the operator

$$L = j \begin{pmatrix} \frac{\partial}{\partial t} & -q(t) \\ sq^*(t) & -\frac{\partial}{\partial t} \end{pmatrix}, \quad (5)$$

where  $q(t)$  is signal. The eigenvalues  $\lambda$  of  $L$  are called *generalized frequencies*. They are complex numbers whose real and imaginary parts have physical significance. Let  $v(\lambda, t) = [v_1, v_2]^T$  be an eigenvector of  $L$  corresponding to  $\lambda$ . Nonlinear Fourier coefficients  $a(t, \lambda)$  and  $b(t, \lambda)$  are components of the normalized eigenvector

$$u(t, \lambda) = \begin{pmatrix} a(t, \lambda) \\ b(t, \lambda) \end{pmatrix},$$

where  $u = [e^{j\lambda t} v_1, e^{-j\lambda t} v_2]^T$ . Define  $a(\lambda) \triangleq a(\infty, \lambda)$  and  $b(\lambda) \triangleq b(\infty, \lambda)$ . The value of the NFT at generalized frequency  $\lambda$ , called the *spectral amplitude*, is  $b(\lambda)/a(\lambda)$ .

If  $\lambda_j$  is in the upper half complex plane  $\mathbb{C}^+$ , it can be shown that  $a(\lambda_j) = 0$ , so that the first term in Taylor series of  $a(\lambda)$  at  $\lambda_j$  vanishes [Part I, Section IV. B]. As a result,  $b/a$  reduces to  $b/a'$ , where prime denotes differentiation. It can be shown that  $a(\lambda)$  is an analytic function of  $\lambda$  in  $\mathbb{C}^+$  [Part I, Lemma 4]. Thus the set of generalized frequencies in  $\mathbb{C}^+$  is discrete;  $j \in \mathcal{N} = \{1, \dots, N\}$ .

**Definition 1** (Nonlinear Fourier Transform). The NFT of signal  $q(t)$  is

$$\text{NFT}(q)(\lambda) = \begin{cases} \frac{b(\lambda)}{a(\lambda)}, & \lambda \in \mathbb{R}, \\ \frac{b(\lambda_j)}{a'(\lambda_j)}, & \lambda_j \in \mathbb{C}^+, \quad j \in \mathcal{N}. \end{cases}$$

□

A crucial property of the NFT is that, in the absence of noise, similar to (4)

$$\text{NFT}(q(t, \mathcal{L}))(\lambda) = H(\lambda, \mathcal{L}) \text{NFT}(q(t, 0))(\lambda),$$

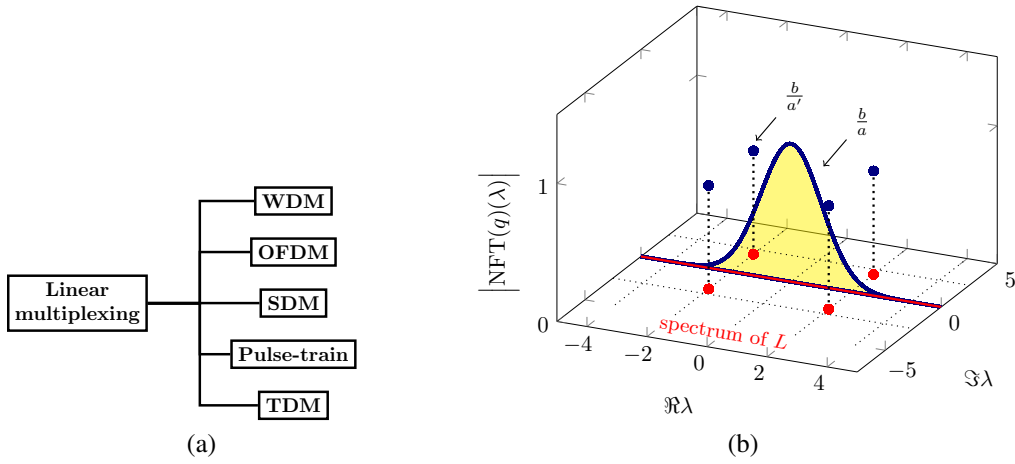


Fig. 3. (a) Linear multiplexing methods. (b) Absolute value of the nonlinear Fourier transform as a 3-dimensional surface on the complex plane.

where  $H$  is the channel filter. For optical fiber, the channel filter is  $H(\lambda) = e^{4js\lambda^2\mathcal{L}}$ . If  $s = 1$ ,  $L$  is self-adjoint (Hermitian); thus  $\mathcal{N}$  is empty and frequencies are real.

In NFDM information is modulated in generalized frequencies  $\lambda$  and their spectral amplitudes  $NFT(\lambda)$ . In the time domain, communication takes place over the nonlinear Fourier basis, instead of complex exponentials  $\exp(j\omega t)$  as in OFDM or WDM. This is a basis made of *nonlinear waves*: Each element of the basis is the inverse NFT of one spectral component, for example a soliton. Superposition of basis elements is nonlinear, described by integral equations [Part I, Eq. 30].

## V. INVERSE NFT

In this section, we present an approach to inverse NFT which is the dual of the forward NFT, and implement it in a continuous- and discrete-time model. The algorithm works in both focusing and defocusing regimes. However, we consider the defocusing regime for simulations.

### A. Inverse NFT

From the eigenvalue equation  $Lv = \lambda v$  for (5),  $q(t)$  can be read off as

$$\begin{aligned} q^*(t) &= s \frac{\partial_t v_2 - j\lambda v_2}{v_1} \\ &= se^{2j\lambda t} \frac{\partial_t b(t, \lambda)}{a(t, \lambda)}. \end{aligned} \quad (6)$$

Equation (6) essentially constitutes the inverse transform.

From the unimodularity condition

$$|a(\lambda)|^2 - s|b(\lambda)|^2 = 1, \quad (7)$$

we obtain

$$|a(\lambda)| = \left( \frac{1}{1 - s|\hat{q}(\lambda)|^2} \right)^{\frac{1}{2}}, \quad (8)$$

where  $\hat{q} = b/a$ .

Since  $a(\lambda)$  is an analytic function of  $\lambda \in \mathbb{C}^+$ , its real and imaginary parts are Hilbert transforms of one another. In the defocusing regime,  $\log a(\lambda)$  is also analytic, thus amplitude

$|a(\lambda)|$  and phase  $\angle a(\lambda)$  are also Hilbert transforms of one another, *i.e.*,

$$\angle a(\lambda) = \mathcal{H}(\log |a(\lambda)|), \quad (9)$$

where  $\mathcal{H}$  denotes Hilbert transform. We refer the reader to Appendix B for proofs.

As a result, we easily obtain  $a$  and  $b = \hat{q}a$  in the defocusing regime. In the focusing regime, integral equations [2, Eq. 30] are solved at one point in time, *e.g.*, at  $t = \infty$ , to obtain  $a(\lambda)$  and  $b(\lambda)$ ,  $\lambda \in \mathbb{R}$ , from  $NFT(q)(\lambda)$ ,  $\lambda \in \mathbb{C}$ . The rest of the algorithms in this section proceed similarly in both regimes.

Having obtained  $a(\lambda)$  and  $b(\lambda)$  for  $\lambda \in \mathbb{R}$ , numerical methods in [Part II], *e.g.*, the layer-peeling (LP) method, can be run inversely in time to obtain  $a(t, \lambda)$  and  $b(t, \lambda)$ . Note that  $a(t, \lambda)$  and  $b(t, \lambda)$  depend on  $\hat{q}(\lambda)$ , but also as on  $q(\tau)$ ,  $\tau > t$ , which is yet unknown. To resolve this, (6) is iterated between  $(a, b)$  and  $q$ : From  $a(t, \lambda)$  and  $b(t, \lambda)$ ,  $q(t)$  is obtained using (6). The resulting  $q(t)$  is then used to obtain  $a(t - \epsilon, \lambda)$  and  $b(t - \epsilon, \lambda)$  using inverse of (6), which is an inverse step in Zakharov-Shabat system (eigenproblem  $Lv = \lambda v$ ), namely (13). This scheme works if  $q(t) \approx q(t - \epsilon)$  as, *e.g.*, in a piecewise constant approximation of  $q(t)$  in Fig. 4 (a).

Equations in the focusing regime in [Part I] can often be generalized by substitutions

$$\begin{aligned} \{a, a^*\} &\rightarrow \{a, a^*\}, & \{q, q^*\} &\rightarrow \{q, -sq^*\}, \\ \{b, b^*\} &\rightarrow \{b, -sb^*\}, & \{\hat{q}, \hat{q}^*\} &\rightarrow \{\hat{q}, -s\hat{q}^*\}, \end{aligned}$$

to include the defocusing regime as well. For example, Parseval's identity for the NFT is

$$E = \int_{-\infty}^{\infty} |q(t)|^2 dt = -\frac{s}{\pi} \int_{-\infty}^{\infty} \log(1 - s|\hat{q}(\lambda)|^2) d\lambda.$$

This implies that  $|\hat{q}(\lambda)| < 1$  in the defocusing regime.

### B. Continuous Layer-Peeling (CLP)

It is preferred to recover  $q(t)$  through an integral representation instead of the differential representation (6). From [Part

I, Eq. 32]

$$q^*(t) = \frac{s}{\pi} \int_{-\infty}^{\infty} \hat{q}(\lambda) e^{2j\lambda t} V_2^1(t, \lambda) d\lambda \quad (10)$$

$$\stackrel{(a)}{=} \frac{s}{\pi} \int_{-\infty}^{\infty} \hat{q}(t, \lambda) e^{2j\lambda t} d\lambda, \quad (11)$$

where  $\hat{q}(t, \lambda) = b(t, \lambda)/a(t, \lambda)$  and  $V^1$  is defined in [Part I, Eq. 27]. Step (a) follows from the causality property in [Part I, Section IV. D]:  $\hat{q}(t, \lambda)$  is in one-to-one relation with  $q(\tau)$  in  $\tau \leq t$ . Thus, in (10) we can assume  $q(\tau) = 0$  for  $\tau > t$ , which gives boundary condition  $V_2^1 = 1$ ; see [Part I, Eq. 38 (4)] and Fig. 4 (a).

For implementation, let  $q(t)$  be supported on  $[T_1, T_2]$ ,  $T = T_2 - T_1$ . Consider a piece-wise constant approximation to  $q(t)$  on the mesh

$$t[k] = T_1 + k\epsilon, \quad \epsilon = T/N, \quad k = 0, \dots, N-1, \quad (12)$$

and set  $q[k] = q(T_1 + k\epsilon)$ .

Nonlinear Fourier coefficients propagate in forward direction according to

$$u[k+1, \lambda] = M[k, \lambda, q] u[k, \lambda], \quad u[0] = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

where the *monodromy* matrix is

$$M \triangleq \begin{pmatrix} x[k] & \bar{y}[k] \\ y[k] & \bar{x}[k] \end{pmatrix}, \quad \det M = 1,$$

in which  $x[k]$  and  $\bar{x}[k]$  are given in [Part II, Eq. 11],

$$y[k] = \frac{sq^*[k]}{\Delta} \sin(\Delta\epsilon) e^{-j\lambda(t[k] + t[k-1])},$$

and  $\bar{y}[k] = sy^*(\lambda^*)$ , where  $\Delta = \sqrt{\lambda^2 - s|q[k]|^2}$ . In the reverse direction

$$u[k, \lambda] = M^{-1}[k, \lambda, q] u[k+1, \lambda], \quad u[N, \lambda] = \begin{pmatrix} a(\lambda) \\ b(\lambda) \end{pmatrix}, \quad (13)$$

where

$$M^{-1}[k, \lambda, q] = \begin{pmatrix} \bar{x}[k] & -\bar{y}[k] \\ -y[k] & x[k] \end{pmatrix}.$$

From the rectangle example in [Part II, Section IV. C] and the piece-wise constant approximation of  $q(t)$ , we see that computing  $a(\lambda)$  requires evaluating  $\cosh(x)$ , where  $x \sim \mu(\{t : \lambda^2 - |q(t)|^2 < 0\})$ , where  $\mu$  denotes Lebesgue measure. Consequently,  $a$  and  $b$  can involve large numbers  $\exp(x)$ ,  $x > 25$ , leading to numerical error — which can further accumulate. We thus update the ratio  $\hat{q} = b/a$  to cancel large numbers

$$\hat{q}[k, \lambda] = \alpha[k] \frac{\bar{\beta}[k] + \hat{q}[k-1, \lambda]}{1 + \beta[k] \hat{q}[k-1, \lambda]}, \quad \hat{q}[0, \lambda] = 0,$$

where

$$\alpha[k] = e^{-2j\lambda(t[k] - t[k-1])} \frac{1 + \frac{j\lambda}{\Delta} \tan(\Delta\epsilon)}{1 - \frac{j\lambda}{\Delta} \tan(\Delta\epsilon)},$$

$$\beta[k] = e^{2j\lambda t[k-1]} \frac{q[k]}{\Delta} \frac{\tan(\Delta\epsilon)}{1 - \frac{j\lambda}{\Delta} \tan(\Delta\epsilon)},$$

and  $\bar{\beta}[k] = s\beta^*[k](\lambda^*)$ . The inverse iteration is

$$\hat{q}[k-1, \lambda] = \frac{\hat{q}[k, \lambda] - \bar{\beta}[k] \alpha[k]}{\alpha[k] - \beta[k] \hat{q}[k, \lambda]}, \quad \hat{q}[N, \lambda] = \hat{q}(\lambda).$$

The discontinuity of  $q(t)$  in the piece-wise constant approximation causes numerical error in the Fourier transform relation (11) due to Gibbs phenomenon. The  $L$  eigenproblem is still well-posed (for appropriate signals), unless  $q(t)$  contains delta functions, in which case  $v(t, \lambda)$  is not differentiable in  $t$ . At discontinuities,  $q(t)$  in (11) is the average of  $q[k] = q(T_1 + (k-1)\epsilon - \delta)$  ( $\delta \rightarrow 0^+$ ) and zero, therefore

$$q[k] = \frac{2s}{\pi} \int_{-\infty}^{\infty} \hat{q}[k, \lambda] e^{2j\lambda t[k]} d\lambda. \quad (14)$$

Integral (14) is evaluated on the mesh

$$\lambda[m] = L_1 + m\mu, \quad \mu = L/M, \quad m = 0, \dots, M-1,$$

where NFT is supported on (simulation) nonlinear band  $[L_1, L_2]$ ,  $L = L_2 - L_1$ .

There are issues with CLP, discussed in Section V-D. As a result, we consider the discrete layer-peeling (DLP) algorithm. Certain discretizations of the Zakharov-Shabat system are captured by  $z = \exp(-2j\lambda\epsilon)$ . Here  $u$  can be represented in terms of powers of  $z$ . A polynomial approach to the Zakharov-Shabat system is discussed in several papers by Wahls and co-authors [9]–[12]. Polynomials make it easier to see places where divide-and-conquer strategies can be devised. Discrete algorithms are also accurate. Coefficients of a polynomial  $p(z)$  evaluated at  $z = \exp(-2j\lambda\epsilon)$  are coefficients of the Fourier series of the  $\pi/\epsilon$ -periodic function  $p(z(\lambda))$ . Thus CLP and DLP can be viewed as frequency and time domain approaches to integrating first-order ordinary differential equations. We refer the reader to [11, Section II] and references therein for the fast DLP. Below, we consider the basic DLP and present derivations and proofs omitted in [11, Section II].

### C. Discrete Layer-Peeling (DLP)

In CLP  $x[k]$  and  $y[k]$  depend on  $\cos(D\epsilon)$  and have to be evaluated for every  $\lambda \in [L_1, L_2]$ . A finite-dimensional representation, where  $u[k, \lambda]$  depends on a finite set of parameters is preferred. Even though hyperbolic functions and  $M$  can be approximated in terms of  $z$  and  $z^{-1}$ , it is equally good and simpler to consider the Ablowitz-Ladik (AL) method [Part II, Eq. 17]

$$v[k+1, z] = c_k \begin{pmatrix} z^{\frac{1}{2}} & Q[k] \\ sQ^*[k] & z^{-\frac{1}{2}} \end{pmatrix} v[k, z], \quad (15)$$

$$v[0, z] = \begin{pmatrix} 1 \\ 0 \end{pmatrix} z^{\frac{k_0}{2}}, \quad 0 \leq k \leq N-1,$$

where  $Q[k] = \epsilon q[k]$ ,  $c_k = 1/\sqrt{1 - s|Q[k]|^2}$ , and  $k_0 = T_1/\epsilon$ . Note that approximation  $1 - 2j\lambda\epsilon \approx \exp(-2j\lambda\epsilon)$  in [Part II, Section III. E] makes  $v(t, \lambda)$  in the AL method a periodic function of  $\lambda$  with period  $L = \pi/\epsilon$ . This is in contrast with CLP where  $L$  is arbitrary.

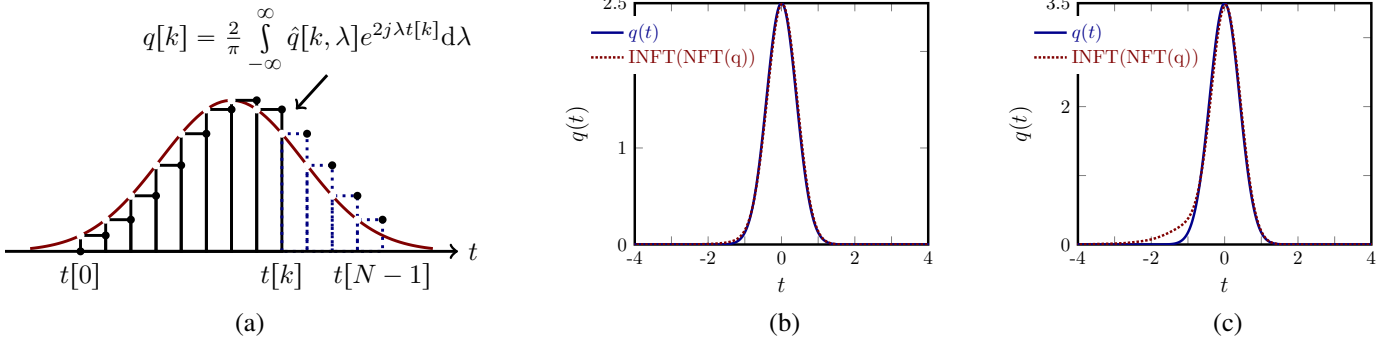


Fig. 4. (a) CLP with piece-wise constant approximation. (b, c) Reconstructions for two amplitudes. Parameters are  $T = L = 64$  and  $N = M = 2048$ .

Discretizing  $\hat{q}(\lambda, t) = \exp(-2j\lambda t)(v_2/v_1)$  using (12)

$$\begin{aligned} \hat{q}[k+1] &= z^{k+k_0} \frac{v_2[k+1]}{v_1[k+1]} \\ &= z^{k+k_0} \frac{sQ^*[k]v_1[k] + z^{-\frac{1}{2}}v_2[k]}{z^{\frac{1}{2}}v_1[k] + Q[k]v_2[k]} \\ &= z^{k+k_0-\frac{1}{2}} \frac{sQ^*[k] + z^{-\frac{1}{2}}(v_2[k]/v_1[k])}{1 + Q[k]z^{-\frac{1}{2}}(v_2[k]/v_1[k])}. \end{aligned}$$

By induction, we can see that  $\deg(v_1) = \deg(v_2) + 1/2$ , where  $\deg(v)$  is the highest positive power of  $z$  in  $v$ . Thus  $\lim_{z \rightarrow \infty} z^{-\frac{1}{2}}v_2[k]/v_1[k] = 0$  and

$$\begin{aligned} Q^*[k] &= s \lim_{z \rightarrow \infty} z^{-k-k_0+\frac{1}{2}} \hat{q}[k+1] \\ &= s \lim_{z \rightarrow \infty} z^{-k-k_0+\frac{1}{2}} \frac{b[k+1, z]}{a[k+1, z]} \\ &= s \frac{B[k+1, \infty]}{A[k+1, \infty]} \\ &= s \frac{B_0[k+1]}{A_0[k+1]}. \end{aligned} \quad (16)$$

Here

$$A[k, z] = \sum_{m=0}^{M-1} A_m[k] z^{-m}, \quad A_0[k] = 1, \quad (17a)$$

$$B[k, z] = \sum_{m=0}^{M-1} B_m[k] z^{-m}, \quad (17b)$$

and

$$a[k, z] = z^{-\frac{k+k_0}{2}} v_1[k, z], \quad b[k, z] = z^{\frac{k_0+k}{2}} v_2[k, z], \quad (18a)$$

$$A[k, z] = a[k, z], \quad B[k, z] = z^{-(k_0+k)+\frac{1}{2}} b[k, z]. \quad (18b)$$

From (15) and (18a)–(18b), we obtain forward DLP iterations

$$\begin{pmatrix} A[k+1] \\ B[k+1] \end{pmatrix} = c_k \begin{pmatrix} 1 & Q[k]z^{-1} \\ sQ^*[k] & z^{-1} \end{pmatrix} \begin{pmatrix} A[k] \\ B[k] \end{pmatrix},$$

$$\begin{pmatrix} A[0] \\ B[0] \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad 0 \leq k \leq N-1,$$

and inverse iterations starting from  $k = N-1$

$$\begin{pmatrix} A[k] \\ B[k] \end{pmatrix} = c_k \begin{pmatrix} 1 & -Q[k] \\ -sQ^*[k]z & z \end{pmatrix} \begin{pmatrix} A[k+1] \\ B[k+1] \end{pmatrix}. \quad (19)$$

Fourier coefficients in (17a)–(17b) are computed as

$$\begin{aligned} A_m[k] &= \frac{1}{L} \int_{L_1}^{L_2} A[k, e^{-2j\lambda t[k]}] e^{-2jm\epsilon\lambda} d\lambda \\ &= e^{-2j\pi \frac{L_1}{L} m} \frac{1}{M} \text{DFT}(A)[m+1], \end{aligned} \quad (20)$$

where  $\text{DFT}(A)$  is the discrete Fourier transform of the samples of  $A[k, e^{-2j\lambda t[k]}]$  in  $\lambda$ .

The unimodularity condition (7) in DLP is discretized to

$$\begin{aligned} A_{k+1}(z)A_{k+1}^*\left(\frac{1}{z^*}\right) \\ -sB_{k+1}(z)B_{k+1}^*\left(\frac{1}{z^*}\right) = \left( \prod_{i=1}^{k+1} \text{sign}(c_i^{-2}) \right), \quad z \in \mathbb{C}. \end{aligned}$$

On the unit circle  $z = 1/z^*$  and

$$|A_k(z)|^2 - s|B_k(z)|^2 = 1, \quad |z| = 1,$$

where (23) is assumed, so that  $c_i > 0$ . In frequency domain

$$\underline{A} * \tilde{\underline{A}} - s\underline{B} * \tilde{\underline{B}} = \delta[m], \quad (21)$$

where  $\delta[m]$  is Kronecker delta function, and  $\tilde{\cdot}$  is the flip operator, i.e.,  $\tilde{\underline{A}}_m = \underline{A}_{M-m+1}$  and  $\tilde{\underline{B}}_m = \underline{B}_{M-m+1}$ . Identity (21) can be checked in iterations to ensure accuracy.

DLP is summarized in Algorithm 1.

#### D. Remarks on LP Algorithms

The left hand side of (6) is independent of  $\lambda$ . Thus, one can set, e.g.,  $\lambda = \infty$  in (6). As a result, (16) can be directly obtained from (6) by substituting  $\exp(2j\lambda t)$  with  $z^{-k-k_0}$  and  $\partial_t$  with  $z^{1/2}/\epsilon$ , and taking the zeroth-order term in the  $z$  expansion of both sides.

The forward LP algorithm is a basic integration scheme in numerical analysis. It applies well to differential equations with time-varying coefficients, which are exactly solvable in the special case that coefficients are constant. For the operator (5), forward LP is discussed in [13] and [Part II]. The reverse iterations are essentially same as forward iterations. Signal recovery in the inverse LP is simply through Zakharov-Shabat system via (6) or its discretization (16). Inverse CLP and DLP are widely used in fiber Bragg grating design; see [14]–[16] and references therein. For transmission application,



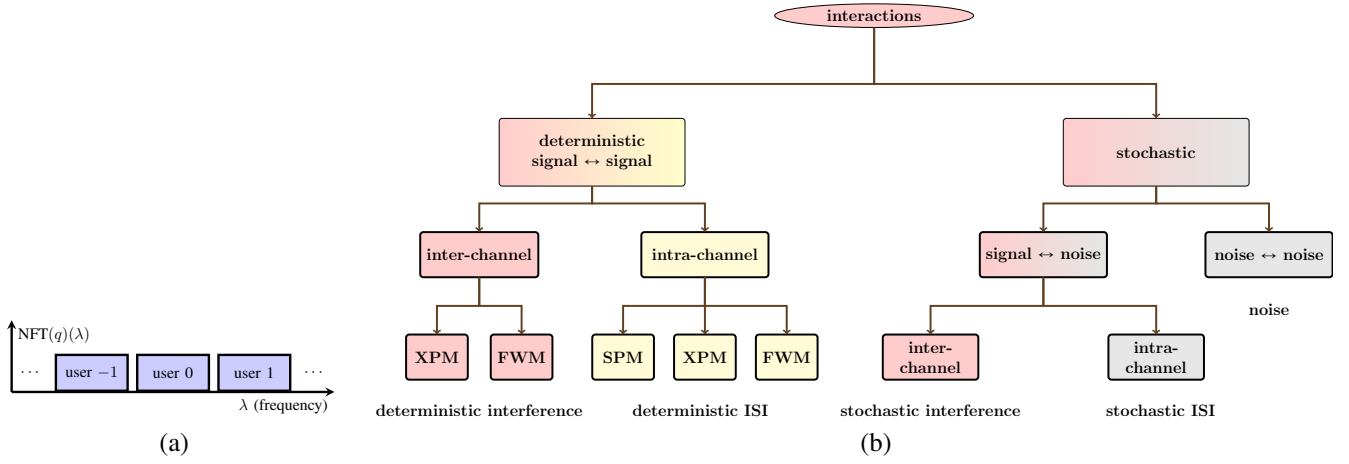


Fig. 5. (a) Multi-user NFD M. (b) Interactions in linear multiplexing. Interactions that cannot be removed are identified with red. Interactions that can *partially* be equalized using signal processing are shown by yellow. Grey terms are addressed by coding. SPM and XPM stand for self- and cross-phase modulation, and FWM is four-wave mixing.

---

**Algorithm 1** Discrete Layer-Peeling Algorithm

---

Use the Hilbert transform to obtain  $a(\lambda)$  and  $b(\lambda)$ , using (8) and (9).

Set  $a[N, \cdot] \triangleq a(\lambda)$  and  $b[N, \cdot] \triangleq b(\lambda)$  and obtain  $A[N, \cdot]$  and  $B[N, \cdot]$  from (18b), as functions of  $\lambda$ .

Compute Fourier coefficients in (17a)–(17b) using (20).

Define vector of Fourier coefficients

$$\underline{A} \triangleq (A_1[N], \dots, A_M[N]),$$

$$\underline{B} \triangleq (B_1[N], \dots, B_M[N]).$$

**for**  $k = N, \dots, 1$  **do**

Obtain  $Q[k]$  from (16) or (34), and  $q[k] = Q[k]/\epsilon$ .

Update

$$A \leftarrow \frac{1}{c_k} (A - Q[k]B),$$

$$B \leftarrow \frac{1}{c_k} \text{shift}(-sQ^*[k]A + B),$$

where  $\text{shift}(x) = (x_2, \dots, x_n, 0)$  is left shift.

**end for**

---

these algorithms should be optimized and tested so that they withstand large amplitudes and signal variations.

In our evaluation, we found that CLP suffers from the following problems:

- 1) Piece-wise constant approximation of  $q(t)$  is discontinuous in time. Consequently, (14) is inexact.
- 2) Reconstruction is sensitive to spectral parameters  $L$  and  $M$  and, to a less extent, temporal variables  $T$  and  $N$ . In particular, integral (14) should be accurately evaluated.
- 3) Local errors resulting from finiteness of  $L$  and  $M$  and Gibbs phenomenon accumulate. The algorithm is divergent at high signal amplitudes; see Figs. 4 (b–c).

In DLP, on the other hand, it should be ensured that  $a(\lambda)$  and  $b(\lambda)$  are represented by sufficient number of Fourier coefficients in (17a)–(17b). Once  $a(\lambda)$  and  $b(\lambda)$  are accurately

represented, the rest of the algorithm is exact. In particular, (16), in contrast to its analogue in CLP, is exact. DLP can be implemented directly in  $\lambda$  domain as well: Iterations (19) can be run for each  $\lambda$ , instead of updating Fourier coefficients. The improved accuracy of DLP seems to stem from exactness of (16), not from the temporal or frequency domain implementations.

Error in DLP can result from large numbers in polynomial coefficients. Unlike CLP, the ratio  $b/a$  is not a polynomial and cannot be similarly updated without approximation. Variables in DLP grow with scale factor  $\exp(P)$ , where

$$P = -\frac{1}{2} \sum_{k=1}^N \log(1 - |Q[k]|^2).$$

As  $N \rightarrow \infty$  for fixed  $T$ ,  $P \rightarrow 0$ . We require that

$$0 \leq P < P_{\max} \quad (22)$$

for a moderate  $P_{\max}$ , and

$$|Q[k]| < 1. \quad (23)$$

It can be shown that

$$\frac{d|Q[k-1]|}{d|Q[k]|} \propto \frac{1}{(1 - s|Q[k]|^2)^2}.$$

It is thus desirable to have

$$|Q[k]| \ll 1, \quad (24)$$

so that  $Q[k-1]$  is not sensitive to error in  $Q[k]$ . The applicability conditions of DLP are (22), (23), and preferably (24) as well.

Note that although  $\det M = 1$ , numerical error could perturb eigenvalues off the unit circle, so that for one eigenvalue  $|\lambda_1(M)| < 1$ , and potentially upon further iterations  $\lambda_1 \rightarrow 0$ . That is to say, CLP and DLP iterations could still be non-invertible for large signals, even as  $\det M = 1$  analytically.

## VI. LINEAR AND NONLINEAR MULTIPLEXING

Nonlinear multiplexing is described in [Part I–III] in general. However, multi-user simulations in [Part III] were combined with WDM. In the defocusing regime, multi-user NFDm simplifies, as shown in this section. We first review linear multiplexing.

### A. Linear Multiplexing

Let  $\mathcal{H}$  be a Hilbert space of signals at the input of a channel, with an orthonormal basis  $\{\phi_k\}$ . Partition  $\mathcal{H} = \cup_{k=1}^{N_u} \mathcal{H}_k$ , where  $\mathcal{H}_k$  is spanned by  $\{\phi_i\}_{i \in I_k}$ ,  $I_k \subset \mathbb{N}$ ,  $I_i \cap I_j = \emptyset$ . Linear multiplexing corresponds to adding the signals of  $N_u$  users in orthogonal subspaces  $\mathcal{H}_k$ ,  $1 \leq k \leq N_u$ . Consider the linear channel (1). Here,  $T : \mathcal{H}_k \mapsto \mathcal{H}'_k$ , where  $\mathcal{H}'_k$  in general are not orthogonal. However, if  $\mathcal{H}$  is partitioned according to the channel eigenvectors, then  $\mathcal{H}'_k = \mathcal{H}_k$  are orthogonal. This results in a diagonal linear multiplexing, where user subspaces are orthogonal from input to output.

### B. Nonlinear Multiplexing

We now consider nonlinear multiplexing for multi-user communication based on the NFT.

1) *Signal Spaces*: Let  $\hat{\mathcal{Q}}$  be the space of signals  $\hat{q}(\lambda) : \mathcal{W} \mapsto \mathbb{D}$ , where  $\mathcal{W} = [-\pi W, \pi W]$  and  $\mathbb{D}$  is the open unit disk, with “norm”

$$\|\hat{q}(\lambda)\|_* = \left( -\frac{1}{\pi} \int_{-\infty}^{\infty} \log(1 - |\hat{q}(\lambda)|^2) d\lambda \right)^{\frac{1}{2}}.$$

In the defocusing regime  $|\hat{q}| < 1$ . We introduce a transformation  $F : \hat{\mathcal{Q}} \mapsto \mathcal{U}$  so as to map the unit disk to the complex plane. An appropriate choice is

$$U(\lambda, z) = \left( -\log(1 - |\hat{q}(\lambda, z)|^2) \right)^{\frac{1}{2}} e^{j\angle \hat{q}(\lambda, z)}, \quad (25)$$

where  $\angle \hat{q}$  is the phase of  $\hat{q} \in \hat{\mathcal{Q}}$  and  $U \in \mathcal{U}$ . The image  $\mathcal{U}$  of  $F$  is the vector space of finite-energy signals supported on  $\mathcal{W}$ . The domain  $\hat{\mathcal{Q}}$  of  $F$  is not a vector space with addition and multiplication, thus unsuitable for a communication theory.

As usual in communication theory of waveform channels [17, Chapter 8], we expand a signal  $U \in \mathcal{U}$  onto an orthonormal basis of  $\mathcal{U}$ . In multi-user NFDm, we partition  $\mathcal{U} = \bigcup_k \mathcal{U}_k$  into user subspaces  $\mathcal{U}_k$ . User  $k$  operates in subspace  $\mathcal{U}_k$  spanned by basis  $\{\Phi_\ell^k(\lambda)\}_{\ell=-\infty}^{\infty}$ . If there are  $N_u$  users, each having  $N_s$  degrees-of-freedom (symbols), then

$$U(\lambda, 0) = \sum_{k=k_1}^{k_2} \sum_{\ell=\ell_1}^{\ell_2} s_\ell^k \Phi_\ell^k(\lambda), \quad (26)$$

where

$$k_1 = -\left\lfloor \frac{N_u}{2} \right\rfloor, \quad k_2 = \left\lceil \frac{N_u}{2} \right\rceil - 1,$$

where  $\lfloor x \rfloor$  and  $\lceil x \rceil$  denote, respectively, rounding  $x \in \mathbb{R}$  to nearest integers towards minus and plus infinity. Similar relations hold for  $\ell_1$  and  $\ell_2$  in terms of  $N_s$ .

In general, user signals can overlap in time or (generalized) frequency. Interference in user  $k$  is in the orthogonal subspace

$\mathcal{U}_k^\perp$  and can be projected out. As a special case, we consider non-overlapping users operating in equally spaced nonlinear bands of width  $W_0$  Hz (in  $l = \lambda/2\pi$ ). User  $k$  is centered at  $l = kW_0$  and operates in the nonlinear band

$$\mathcal{W}_k = \left[ kW_0 - \frac{W_0}{2}, kW_0 + \frac{W_0}{2} \right], \quad k_1 \leq k \leq k_2.$$

In this case,  $\Phi_\ell^k(\lambda) = \Phi_\ell(\lambda - 2\pi kW_0)$  and  $W = N_u W_0$ .

Taking the inverse Fourier transform  $\mathcal{F}^{-1}$  of (26)

$$u(\tau, z) = \sqrt{2} \mathcal{F}^{-1}(U(\lambda, z)) \quad (27)$$

$$= \sqrt{2} \sum_{k=k_1}^{k_2} \left( \sum_{\ell=\ell_1}^{\ell_2} s_\ell^k \phi_\ell(\tau) \right) e^{j2\pi kW_0 \tau}. \quad (28)$$

Factor  $\sqrt{2}$  is introduced to balance energies in Section VI-B2. Variable  $\tau$  can be interpreted as a *generalized time* (measured in seconds), in Fourier transform relation with generalized frequency  $l = \lambda/2\pi$  (measured in Hz). It coincides with physical time  $2t$  for small amplitude signals  $\|q\|_{L^1} \ll 1$ . Typically,  $\phi_\ell(\tau) = \phi(\tau - \ell T_0)$ ,  $T_0 = 1/W_0$ , where  $\phi(\tau)$  satisfies the Nyquist criterion for zero-ISI. NFDm signal (28) is identical to the WDM signal; see [Part III, Eq. 2]. It simply represents every sample of  $U(\lambda, z)$  as a degree-of-freedom.

2) *NFDm Transmitter*: Modulation begins with selecting matrix of symbols  $[s_\ell^k]$  according to a constellation  $\Xi$  and computing (26). This can be done directly in the spectral domain (26), or starting in the temporal domain (28). The transmitted signal in the spectral domain is

$$\hat{q}(\lambda, 0) = \left( 1 - e^{-|U(\lambda, 0)|^2} \right)^{\frac{1}{2}} e^{j\angle U(\lambda, 0)}, \quad (29)$$

where  $U(\lambda, 0) = \mathcal{F}(u(\tau, 0))/\sqrt{2}$ . Finally,

$$q(t, 0) = \text{INFT}(\hat{q}(\lambda, 0)).$$

The (linear) Fourier spectrum  $Q(f, 0) = \mathcal{F}(q(t, 0))$  has bandwidth  $B$  Hz. When  $\|q(t, 0)\|_{L^1} \ll 1$ ,  $q(t, 0) \approx -\sqrt{2}u^*(2t, 0)$  and  $W \approx 2B$ .

Equations (25) and (29) ensure that

$$|U(\lambda, 0)|^2 = -\log(1 - |\hat{q}(\lambda, 0)|^2),$$

so that  $[s_\ell^k]$ ,  $\hat{q}(\lambda, 0)$  and  $q(t, 0)$  have the same energies

$$\begin{aligned} \sum_{l,k} |s_\ell^k|^2 &= -\frac{1}{\pi} \int_{-\infty}^{\infty} \log(1 - |\hat{q}(\lambda, 0)|^2) d\lambda \\ &= \int_{-\infty}^{\infty} |q(t, 0)|^2 dt. \end{aligned}$$

3) *NFDm Receiver*: At the receiver, after NFT,  $U(\lambda, \mathcal{L})$  is computed using (25). The received symbol is

$$\hat{s}_\ell^k = \int_{-\infty}^{\infty} U(\lambda, \mathcal{L}) \Phi_\ell^k(\lambda) d\lambda. \quad (30)$$

Alternatively,  $u(\tau, \mathcal{L})$  can be computed using (27). The received symbol at the output is obtained by match filtering

$$\hat{s}_\ell^k = \int_{-\infty}^{\infty} u(\tau, \mathcal{L}) \phi_\ell^k(\tau) d\tau.$$



*Remark 1.* It is not necessary to filter out-of-band signals and noise, as projection onto basis functions in (30) does that.  $\square$

4) *Uniform and Geometric Constellations:* We choose a multi-ring constellation  $\Xi$  for symbols  $s_\ell^k$  in  $\tau$  domain. There are  $N_r$  uniformly-spaced rings with radii in interval  $[a, b]$  and ring spacing  $\Delta r$ , and  $N_\phi$  uniformly spaced phases. For  $N_s = 1$  and  $a = 0$ , this leads to a geometric constellation  $\Xi'$  in  $\lambda$  domain with phases in  $\Xi$  and radii

$$r_n^2 = 1 - e^{-c(n-1)^2}, \quad c = \frac{1}{2}(\Delta r)^2. \quad (31)$$

Constellations are shown in Fig. 10(a)–(b). As  $r_n \rightarrow 1$ , the distance between rings in (31) decreases exponentially. In this way, an infinite number of choices is realized in the finite interval  $0 \leq |\hat{q}| < 1$ , such that signal energies are uniformly separated.

## VII. NFDM AND WDM ACHIEVABLE RATES

In this section, we compare NFDM and WDM data rates, with the same bandwidth and power constraints. We demonstrate the impact of interference and ISI in both methods and their spectral efficiencies.

We first illustrate the main ideas by a simple simulation, without bothering with insignificant details. Let  $N_u = 15$  and  $\phi(\tau)$  be a root raised-cosine function with an excess bandwidth factor denoted by  $r$ . We consider a 15-user NFDM and a 15-user WDM with the same overall bandwidth  $B = 60$  GHz and power  $\mathcal{P} = E/T$ , where  $E$  and  $T$  are energy and time duration of the (entire) multiplexed signal  $q(t, 0)$ . In this section, time duration and bandwidth are defined as intervals containing 99% of the signal energy [Part III]. The per-user bandwidth in 15-user WDM is 4 GHz. In NFDM, users can overlap in (linear) frequency and use the whole bandwidth. In this illustrative simulation, noise is set to zero. Fiber parameters are given in [Part III, Table I], with dispersion  $D = -17$  ps/(nm·km). For numerical computations, a simulation nonlinear bandwidth  $L > W$  is chosen so that  $\epsilon$  is small enough.

Figs. 8(a)–(b) show a sample 15-user WDM and NFDM transmit signals at power  $\mathcal{P} = 0$  dBm. These two signals take the same bandwidth  $B = 60$  GHz and time duration 1.5 ns. The corresponding received signals at  $\mathcal{L} = 2000$  km, with one “ADM” at the receiver and in the absence of noise, are shown in Figs. 7(a)–(b). It can be seen that WDM users interfere with another at this power, while NFDM symbols can be perfectly recovered. In WDM, part of distortion in Fig. 7(b) can be removed by back-propagation. The resulting mismatch between signals in Fig. 7(c) cannot be compensated for in a network environment; we refer to this distortion as *data rate bottleneck in linear multiplexing*. This distortion is increased as  $N_s$  or number of ADMs is increased.

Denote the *matrix* of transmit and received symbols by  $S = [s_\ell^k]$  and  $\hat{S} = [\hat{s}_\ell^k]$ . We tested transmission of random symbols at high powers  $\mathcal{P} \leq 3$  dBm in NFDM and recovered the *matrix* of transmit symbols nearly perfectly for various  $N_u \leq 15$  and  $N_s \leq 15$ . The relative error  $e = \|\hat{S} - S\| / \|S\| \approx 10^{-3} - 10^{-9}$ , arising from numerical errors, is quite small. This supports that, in the absence of noise, large spectral efficiencies

may be achieved in NFDM. In principle, the achievable rate is infinite in deterministic bandlimited model: Infinite number of degrees-of-freedom exists in any finite bandwidth, *e.g.*, in (26), which can be perfectly recovered. Spectral broadening approaches zero as input power tends to zero, or is finite at finite powers, leading to infinite spectral efficiency. In contrast, the corresponding rate of WDM with the same constraints is small (approaching zero with transmit power), as shown in Fig. 2. Note that noise does not substantially influence WDM rates at high powers in Fig. 2.

In the second simulation, we introduce distributed noise and conduct several numerical studies. Constellation  $\Xi$  consists of  $N_r = 32$  rings each with  $N_\phi = 128$  phase points, with  $a = 0.7$  and  $b = 1.6$ . We choose  $a > 0$  to operate at high power. The value of  $b$  is twice of that in Fig. 7, so that on average a behavior similar to Fig 7 is obtained. The excess bandwidth in the root raised-cosine function is  $r = 25\%$ . Noise power spectral density  $\sigma_0^2$  is based on realistic parameters in [Part III, Table I]. We set  $N_u = 15$ , while  $N_s = 1$  so that  $N$  is not too large for several thousand simulations. The highest energy signal is the one with  $s_\ell^k = b$ ,  $\forall k, \ell$ . We fix noise bandwidth for the entire simulation to be the maximum input output bandwidth of the highest energy signal. The number of signal samples in time and generalized frequency is  $N = M = 16384$ . One nonlinear de-multiplexer (a filter) is implemented at  $\mathcal{L} = 2000$  km, followed by channel inversion  $H^*(\lambda, \mathcal{L})$  in the nonlinear Fourier domain. We calculate transition probabilities  $s_0^0 \mapsto \hat{s}_0^0$  based on 4000 simulations of the stochastic NLS equation. Data rate in the memoryless model is then numerically estimated by maximizing mutual information subject to average power constraint. We obtain a data rate

$$R_{\text{NFDM}} = 10.5, \quad \text{bits/2D}, \quad (32)$$

at (total) average input power  $\mathcal{P} = -0.33$  dBm, where bits/2D is bits per two real (one complex) degrees-of-freedom.

Received symbols corresponding to four transmit symbols  $s_0^0 = 0.7$ ,  $s_0^0 = 0.98$ ,  $s_0^0 = 1.2624$ ,  $s_0^0 = 1.6$  are shown in Fig. 10 (c). The size of “clouds” does not increase as  $|s_0^0|$  is increased.

We now consider a 15-user WDM system with the same constraints as in NFDM, namely, 60 GHz bandwidth and  $\mathcal{P} = -0.33$  dBm average input power. A constellation with  $N_r = 32$  uniformly spaced rings, with  $a = 0.73$  and  $b = 1.45$ , and with  $N_\phi = 128$  points on each ring is used. The excess bandwidth in the root raised-cosine function is  $r = 25\%$ , and  $N_s = 1$ . One linear ADM (a filter) is implemented at  $\mathcal{L} = 2000$  km, followed by back-propagation. Transition probabilities  $s_0^0 \mapsto \hat{s}_0^0$  are computed based on 4000 channel simulations. We obtain data rate

$$R_{\text{WDM}} = 5.26, \quad \text{bits/2D}. \quad (33)$$

Fig. 11 shows received symbols for four transmit symbols  $s_0^0 = 0.7527$ ,  $s_0^0 = 0.9575$ ,  $s_0^0 = 1.185$ ,  $s_0^0 = 1.458$  in WDM. WDM “clouds” seem bigger than the corresponding ones in NFDM at the same power in Fig. 10 (c). Note that in WDM, there is a constant rotation of symbols, even after back-propagation. The rotation, which is about  $\gamma \mathcal{L} \mathcal{P}$  ( $\gamma$  being

nonlinearity coefficient), can be explained from [8, Eq. 21]. Note also that received symbols associated with  $s_0^0 = 1.485$  cluster around a ring with radius smaller than 1.458. This can be explained via Fig. 7 (b): Signal energy for the central user  $k = 0$  at the output is lower than the corresponding energy at the input, because of the spectral broadening. Back-propagation is energy-preserving, giving rise to a received symbol with reduced absolute value. This is in contrast with users  $k = \pm 7$  on the boundaries of spectral band. As the input amplitude tends to infinity, all energy could transfer to out of band, as a result of the so-called *forward cascade* in (weak) wave turbulence theory [8, Section V]. Spectral broadening occurs in NFDM too, however, in NFDM users are not filtered; they can interact with one another in time and in (ordinary) frequency. The overall spectral broadening beyond 60 GHz is small, and is a little higher in WDM than in NFDM. This could be because WDM spectrum varies rapidly; note the deep wells in the input signal in Fig. 7 (b).

For comparison, the capacity of an AWGN channel with signal-to-noise ratio  $\text{SNR} = \mathcal{P}/(\sigma_0^2 B \mathcal{L})$  is

$$\begin{aligned} C_{\text{AWGN}} &= \log_2(1 + \text{SNR}) \\ &= 13.78, \quad \text{bits/2D}. \end{aligned}$$

Note that time duration in our simulations is according to the 99%-energy definition, which is higher than  $T_s = 1/B_0$ ,  $B_0 = B/N_u$ , that is often assumed in digital communication (e.g., in WDM). Let  $T_k(z)$  and  $W_k(z)$  be 99%-energy time duration and (maximum input output) bandwidth of signal number  $k$  that is sent over channel, at distance  $z$ . In NFDM

$$\begin{aligned} \alpha_{\text{NFDM}} &= \frac{1}{N_u N_s} (ET_k(0)) \times \max_k \max(W_k(0), W_k(\mathcal{L})) \\ &= 6.78. \end{aligned}$$

The corresponding factor in the 15-user WDM is  $\alpha_{\text{WDM}} = 7.6$ . Since  $T_k(0) > T_s$ , using 99%-energy time duration scales down WDM spectral efficiencies in Fig. 2. To compare with data rates in Fig. 2 in bits/s/Hz, one approach is to scale NFDM spectral efficiency

$$R'_{\text{NFDM}} = \frac{\alpha_{\text{WDM}}}{\alpha_{\text{NFDM}}} 10.5 = 11.77, \quad \text{bits/s/Hz}.$$

If 99% time duration is used, spectral efficiencies are

$$\begin{aligned} R'_{\text{NFDM}} &= \frac{10.5}{\alpha_{\text{NFDM}}} = 1.54, \quad \text{bits/s/Hz}, \\ R'_{\text{WDM}} &= \frac{5.26}{\alpha_{\text{WDM}}} = 0.69, \quad \text{bits/s/Hz}. \end{aligned}$$

We point out, however, that the 99%-energy time duration is not an appropriate definition to use in the calculation of the spectral efficiency. The appropriate quantity is bits/2D in a given bandwidth, which is (32) and (33). However, to average over dimensions,  $N_s$  and  $N_u \rightarrow \infty$ .

Fig. 6 shows achievable rates of NFDM and WDM, as well as the capacity of the corresponding AWGN channel. The parameters are  $B = 60$  GHz,  $N_u = 15$ ,  $N_s = 1$ ,  $a = 0$ ,  $32 \leq N_r \leq 64$  and  $N_\phi = 128$ . The power in the horizontal axis is the total power, calculated according to the 99%-energy definition. Data rates, in bits/2D, are computed at optimal input

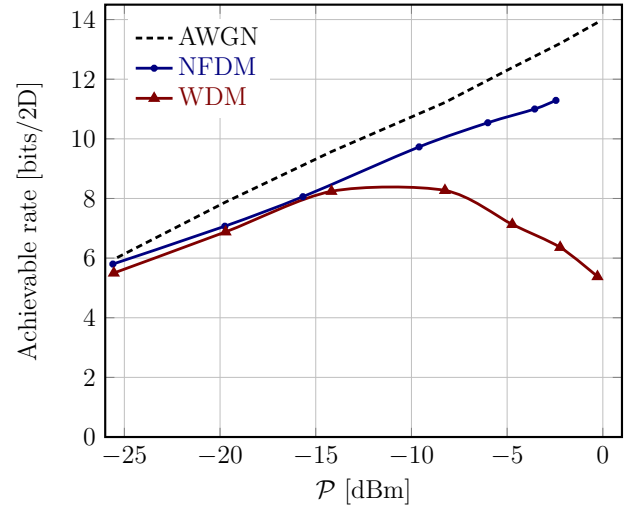


Fig. 6. Achievable rates of the NFDM and WDM, and the capacity of the corresponding AWGN channel.

distribution, obtained using the Arimoto-Blahut algorithm. The rate of the AWGN channel is calculated based on the time duration  $T_s = 1/B_0$ . It can be seen that the NFDM rate increases monotonically with power, in contrast to the WDM rate which characteristically vanishes as the input power is increased more than an optimal value.

Fig. 9(a) shows that the conditional entropy in WDM increases with transmit power, while it is nearly constant in NFDM. This, together with Fig. 9 (b) showing that the conditional probability distribution is shifted with input, could indicate that the channel in the nonlinear Fourier domain behaves approximately as an AWGN channel.

**Remark 2 (Commercial WDM).** The per-user WDM bandwidth is 4 GHz in our simulations, which is smaller than that in commercial systems. Simulations with higher per-user bandwidth, e.g., 40 GHz, need more resources and time. We leave a comprehensive comparison of NFDM and WDM to future work.  $\square$

**Remark 3 (Peak-to-Average Power Ratio).** Similar to OFDM, the peak-to-average ratio of  $U(\lambda, 0)$  could increase as  $N_s \rightarrow \infty$ . This may lead to numerical errors as  $\hat{q}(\lambda) \rightarrow 1$ . Better results may be obtained if data is directly modulated in (26), or other transformations are considered in place of (25).  $\square$

## VIII. COMMENTS ON DISCRETE AND CONTINUOUS SPECTRUM MODULATION

There are several ways that NFT can be used in communications. We review them here, so as to clarify the way that NFT could usefully be applied to data transmission.

### A. Multi-Soliton Communication

Inverse scattering transform is a soliton theory. Fundamental soliton communication was developed in 1980s and 1990s. ON-OFF keying soliton communication was generalized in [18] so as to transmit more than one bit per fundamental soliton. The observation was that eigenvalues are *conserved*,

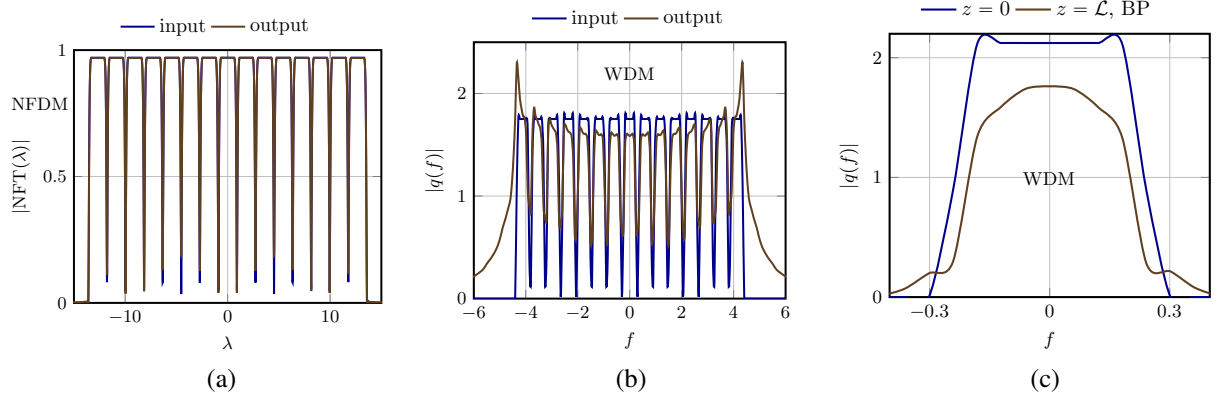


Fig. 7. Deterministic interference in two 15-user methods with the same power, bandwidth and time constraints. (a) Perfect signal reconstruction in NFDM. (b) Inter-channel interference in WDM (a). (c) Interference in WDM after back-propagation.

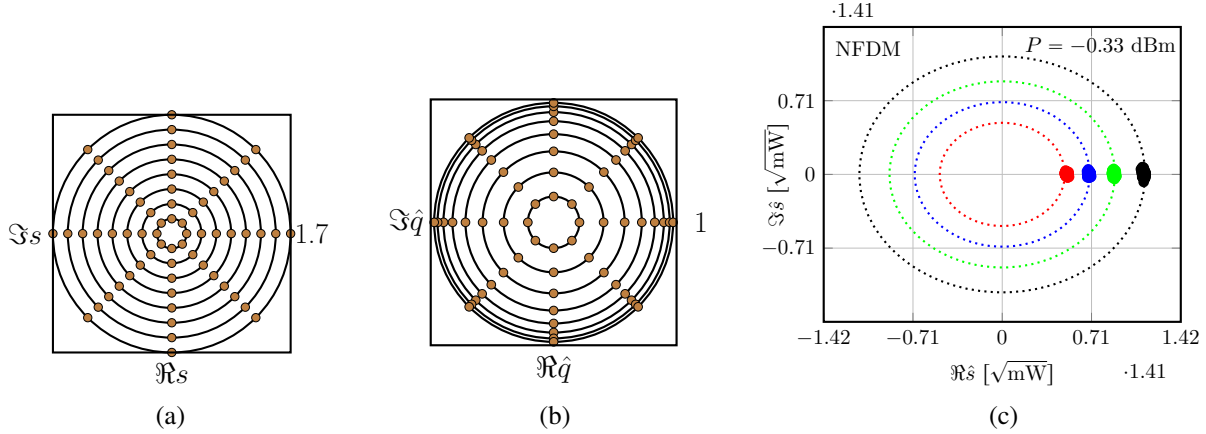


Fig. 10. (a) Uniform constellation in  $\tau$  domain. (b) Geometric constellation in  $\lambda$  domain. (c) Received symbols for four transmit symbols in NFDM.

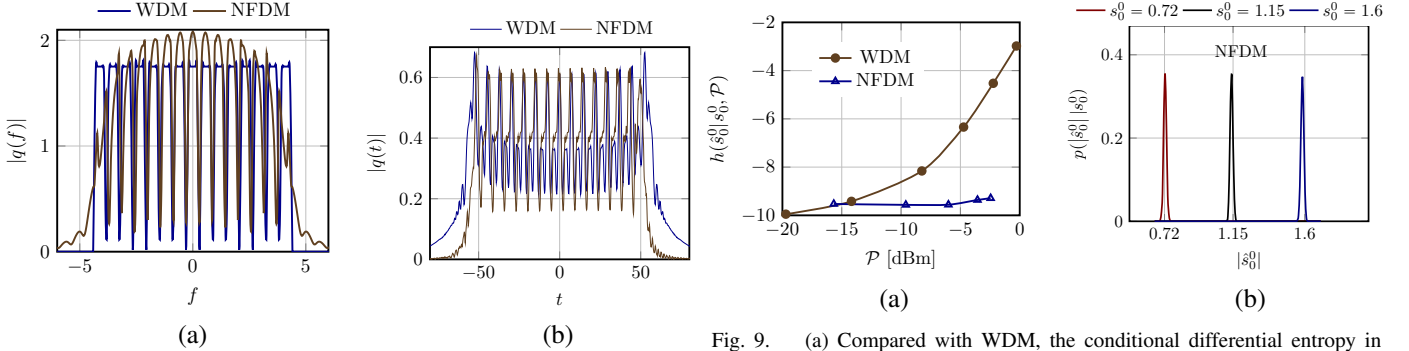


Fig. 8. Bandwidth and time comparison between the proposed 15-user WDM and NFDM at the maximum power  $\mathcal{P} = 0$  dBm. (a) Signals in frequency at  $z = 0$ . Bandwidths are  $B = 60$  GHz. (b) Signals in time at  $z = 2000$  km. Time duration is 1.5 ns.

while amplitude and phase change during the evolution. It is thus natural to put information in invariant quantities, such as eigenvalues. This approach can be extended to general  $N$ -soliton communication.

Data rates of  $N$ -soliton communication are higher than data rates of 1-soliton communication, but roughly equal to data rates of any other  $N$ -bits signal set, for instance, Nyquist signals modulated based on an  $N$ -bits QAM constellation [6]. Evolution of waveforms is not a limitation in communications, since it is equalized at the receiver (as in radio channels).

Fig. 9. (a) Compared with WDM, the conditional differential entropy in NFDM is nearly constant as  $\mathcal{P}$  is increased ( $s_0^0 = 0.2$ ). (b) Conditional probability distribution of  $|\hat{s}_0^0|$  given  $s_0^0$  in NFDM.

Generally, changing pulse shapes does not change data rates (other than minor optimization).

### B. Discrete and Continuous Spectrum Modulation

Degrees-of-freedom in time and nonlinear Fourier domain are in one-to-one correspondence. Thus, modulating both discrete and continuous spectrum does not achieve a data rate higher than existing WDM rates. Indeed, examining WDM signals with vanishing rates in the literature [6], both spectra are (indirectly) fully modulated within bandwidth and power constraint. An approach is not improved by replacing its set

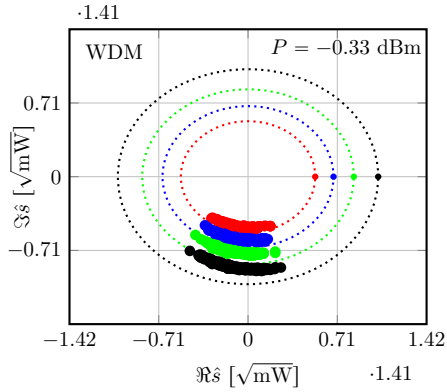


Fig. 11. Received symbols for four transmit symbols in WDM.

of degrees-of-freedom with another equivalent set.

### C. NFT for Signal Multiplexing

NFT can be applied for *signal multiplexing* in multi-user channels. In [Part I–III], a nonlinear frequency-division multiplexing was designed for multi-user networks, in order to address the recent capacity crunch problem arising from interference and ISI in linear multiplexing. In this approach, ADMs will eventually be replaced with nonlinear multiplexers. By using NFT to construct *nonlinear multiplexers*, higher data rates could be achieved.

It is important to note that NFDM is not soliton communication, or discrete and continuous spectrum modulation. Modulating both discrete and continuous spectrum and performing linear multiplexing will not achieve a data rate higher than existing WDM rates [6]. On the other hand, modulating users signals in time domain, and performing nonlinear multiplexing, can improve WDM data rates. NFDM is an approach where users' information is encoded in the nonlinear Fourier domain (non-critical step) and multiplexed nonlinearly (critical step).

## IX. RELATED WORK

Data transmission using the NFT has been studied in several recent closely-related works [2]–[4], [9]–[12], [19]–[30]. In this section, we review some of these papers, not discussed earlier.

### A. Review

Fast nonlinear Fourier transforms were introduced by Wahls et al. in [10]–[12]. In addition to computation speed, discrete algorithms presented in these papers are robust. Discrete algorithms motivated the present paper and enabled nonlinear multiplexing in Section VII.

In [19]–[21], Prilepsky, Le, Turitsyn and co-authors, proposed nonlinear inverse synthesis based on NFT. They demonstrated spectrum modulation and reported low bit error rates [23].

Noise in the nonlinear Fourier domain is studied in [27], [31]–[33]. Distribution of the spectral parameters for the special case that channel is noise-free and input is a white

Gaussian stochastic process is calculated in [31]. A lower bound on spectral efficiency of multi-solitons is obtained in [32]. Noise models in the spectral domain, including models for spectral amplitudes, are studied in [27], [33].

Transmission based on the nonlinear Fourier transform was recently demonstrated in experiments [24]–[26], [28], [30], [34]. Bülow [24]–[26], and subsequently Aref and others [30], showed that one can recover information via eigenvalues and other parts of the nonlinear spectrum. Dong et al. successfully demonstrated error-free transmission of three-eigenvalue NFDM signals over 1800 km [28]. An important consequence of these works is that, the end-to-end channel appears nearly integrable in practice as indicated by the stability of eigenvalues.

### B. Note on Data Rates in Literature

For clarity, we note that, when calculating data rates, one should pay attention to the bandwidth and time duration of signal sets that are used (at input and output). These factors sometimes have not been (fully) taken into account in the past. For example, it certainly does not make sense to compare single-user spectral efficiencies with WDM spectral efficiencies. Furthermore, capacity in bits per “symbol” does not have much implication in communication. Some trivial schemes for fiber channel achieve data rates in bits per symbol growing logarithmically with input power.

## APPENDIX A

### A AND B IN THE ABLOWITZ-LADIK SCHEME

In general,  $A$  and  $B$  are expanded in the basis

$$m = (1 \quad z^{-1} \quad \dots \quad z^{-M+1})^T.$$

The first few iterations of AL scheme are

$$\begin{aligned} \begin{pmatrix} A[0] \\ B[0] \end{pmatrix} &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} A[1] \\ B[1] \end{pmatrix} = c_0 \begin{pmatrix} 1 \\ -Q_0^* \end{pmatrix}, \\ \begin{pmatrix} A[2] \\ B[2] \end{pmatrix} &= c_0 c_1 \begin{pmatrix} 1 - Q_0^* Q_1 z^{-1} \\ -Q_1^* - Q_0^* z^{-1} \end{pmatrix}, \\ \begin{pmatrix} A[3] \\ B[3] \end{pmatrix} &= c_0 c_1 c_2 \begin{pmatrix} 1 - (Q_0^* Q_0 + Q_1^* Q_2) z^{-1} - Q_0^* Q_2 z^{-2} \\ -Q_2^* + (Q_0^* Q_1 Q_2^* - Q_1^*) z^{-1} - Q_0^* z^{-2} \end{pmatrix}, \end{aligned}$$

where  $Q_k \triangleq Q[k]$ . In addition to (6), the following identity can be proved by induction

$$Q[k-1] = \frac{A_{M-1}[k]}{B_{M-1}[k]}. \quad (34)$$

## APPENDIX B

### KRAMERS-KRONIG RELATIONS

Given  $\text{NFT}(q)(\lambda)$ , Riemann-Hilbert integral equations in [Part I, Eq. 30] can be solved at  $t = T$  to obtain  $V^1(T, \lambda) = [a(\lambda), b(\lambda)]$ . This requires solving one system of linear equations. However, there is a simpler method, based on the analyticity of  $a(\lambda)$ .

The real and imaginary parts of a complex analytic function are related via Cauchy-Riemann equations, or their integral analogue, Kramers-Kronig relations.

**Lemma 1** (Kramers-Kronig Relations). *Let  $f(z) = u(x, y) + jv(x, y)$  be a function of a complex variable  $z = x + jy$ , where  $u$  and  $v$  are real-valued functions. Suppose that  $f(z)$  is analytic in  $\mathbb{C}^+$  and  $f(z) \rightarrow 0$  as  $|z| \rightarrow \infty$ . Then*

$$u(x, y) = \mathcal{H}_x(v(x, y)), \quad v(x, y) = -\mathcal{H}_x(u(x, y)). \quad (35)$$

Notation  $\mathcal{H}_x$  denotes Hilbert transform with respect to  $x$

$$\begin{aligned} \mathcal{H}_x(f(x, y)) &= \frac{1}{\pi} \text{p.v.} \int_{-\infty}^{\infty} \frac{f(x', y)}{x' - x} dx' \\ &= \frac{1}{\pi x} * f(x, y), \end{aligned}$$

where  $*_x$  is convolution with respect to  $x$ , and p.v. is principal value.

*Proof:* Consider a point  $z = x + jy \in \mathbb{C}^+$  in the complex plane  $\zeta = x' + jy'$ . Let  $\mathcal{C}$  be a counter-clockwise semi-circular contour, with diameter at  $y' = y$ , and passing slightly around  $z$  from above. We apply Sokhotski-Plemelj formula [Part I, Lemma 8] to  $f(z)$  in  $\mathcal{C}$ . Since  $f(z)$  is analytic inside  $\mathcal{C}$ , 1)  $f(z)$  is differentiable inside  $\mathcal{C}$ , thus satisfies a Hölder condition there with  $\alpha = 1$ ; 2)  $F^-(\zeta) = 0$  in [Part I, Eq. 37]. Thus

$$\begin{aligned} f(x + jy) &= \frac{1}{\pi j} \text{p.v.} \int_{\mathcal{C}} \frac{f(x' + jy)}{x' - x} dx' \\ &\stackrel{(a)}{=} \frac{1}{\pi j} \text{p.v.} \int_{-\infty}^{\infty} \frac{f(x' + jy)}{x' - x} dx'. \end{aligned}$$

Step (a) follows because  $f$  vanishes on arc of  $\mathcal{C}$ . Equating real and imaginary parts of the above equation gives (35). ■

**Corollary 2.** *Under the assumptions of Lemma 1 on  $a(\lambda)$ ,*

$$\angle a(\lambda) = \mathcal{H}(\log |a(\lambda)|). \quad (36)$$

□

*Proof:* From unimodularity condition (7) in defocusing regime

$$\begin{aligned} |a|^2 &= 1 + |b|^2 \\ &\geq 1, \quad \lambda \in \mathbb{R}. \end{aligned}$$

Therefore  $\log |a| \geq 0$  for  $\lambda \in \mathbb{R}$ . We prove  $\log |a| \geq 0$  for  $\lambda \in \mathbb{C}^+$ .

We next analytically extend  $a(\lambda)$  to  $\lambda \in \mathbb{C}^+$  [Part I, Lemma 4]. From [Part I, Eq. 45],

$$a(\lambda) = 1 + \frac{E}{2j\lambda} + \mathcal{O}\left(\frac{1}{\lambda^2}\right), \quad \text{as } |\lambda| \rightarrow \infty,$$

where  $E = \int_{-\infty}^{\infty} |q(t)|^2 dt$ . Thus

$$\begin{aligned} \log a(\lambda) &\rightarrow \frac{E}{2j\lambda} + \mathcal{O}\left(\frac{1}{|\lambda|^2}\right) \\ &\rightarrow 0, \end{aligned}$$

on the arc  $|\lambda| \rightarrow \infty$ . It follows that  $\log |a| \geq 0$  on the boundary of  $\mathbb{C}^+$ . Since  $a(\lambda)$  is analytic in  $\mathbb{C}^+$ ,  $\log |a(\lambda)|$  is a harmonic

function there. From the Maximum Principle,  $\log |a(\lambda)| \geq 0$ , everywhere in  $\mathbb{C}^+$ , thus

$$|a(\lambda)| \geq 1, \quad \forall \lambda \in \mathbb{C}^+.$$

In particular, in the defocusing regime,  $a(\lambda) \neq 0$  in  $\mathbb{C}^+$ .

Using Faa di Bruno's formula, derivatives of  $f(\lambda) = \log a(\lambda)$  can be expressed in terms of derivatives of  $a(\lambda)$ . Since  $a$  is analytic in  $\mathbb{C}^+$ , and  $a \neq 0$ ,  $f(\lambda) = \log a(\lambda)$  is also analytic in  $\mathbb{C}^+$ . Furthermore, since  $f(\lambda)$  is differentiable, it satisfies a Hölder condition with  $\alpha = 1$ ; see [Part I, Lemma 8]. Applying Lemma 1 to  $f(\lambda)$  we obtain (36) in  $\mathbb{C}^+$ . In particular, (36) holds along the line  $\lambda = x + j\epsilon$ ,  $\epsilon > 0$ . Letting  $\epsilon \rightarrow 0^+$  and using continuity of  $a(\lambda)$ , we obtain (36) on the real line. ■

## ACKNOWLEDGMENTS

Financial support of the Institute for Advanced Study at the Technical University of Munich, funded by the German Excellence Initiative, as well as Alexander von Humboldt Foundation, funded by the German Federal Ministry of Education and Research, is acknowledged. The authors thank Frank R. Kschischang and Gerhard Kramer for comments.

## REFERENCES

- [1] X. Yangzhang and M. I. Yousefi, "Nonlinear Fourier transform at defocusing regime," in *2015 Munich Workshop on Information Theory of Optical Fiber*, Dec. 2015, poster presentation. [Online]. Available: [https://www.lnt.ei.tum.de/fileadmin/lnt/events/MIO2015/mio2015poster\\_Yangzhang\\_Xianhe.pdf](https://www.lnt.ei.tum.de/fileadmin/lnt/events/MIO2015/mio2015poster_Yangzhang_Xianhe.pdf)
- [2] M. I. Yousefi and F. R. Kschischang, "Information transmission using the nonlinear Fourier transform, Part I: Mathematical tools," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4312–4328, Jul. 2014. Also published at arXiv, Feb. 2012. [Online]. Available: <http://arxiv.org/abs/1202.3653>
- [3] —, "Information transmission using the nonlinear Fourier transform, Part II: Numerical methods," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4329–4345, Jul. 2014. Also published at arXiv, Apr. 2012. [Online]. Available: <http://arxiv.org/abs/1204.0830>
- [4] —, "Information transmission using the nonlinear Fourier transform, Part III: Spectrum modulation," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4346–4369, Jul. 2014. Also published at arXiv, Feb. 2013. [Online]. Available: <http://arxiv.org/abs/1302.2875>
- [5] M. I. Yousefi, "Information transmission using the nonlinear Fourier transform," Ph.D. dissertation, University of Toronto, Nov. 2012.
- [6] R. J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *IEEE J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 2010.
- [7] M. I. Yousefi and F. R. Kschischang, "On the per-sample capacity of nondispersive optical fibers," *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7522–7541, Nov. 2011.
- [8] M. I. Yousefi, F. R. Kschischang, and G. Kramer, "Kolmogorov-Zakharov model for optical fiber communications," *ArXiv Preprints*, Dec. 2014, arXiv:1411.6550. [Online]. Available: <http://arxiv.org/abs/1411.6550>
- [9] S. Wahls and H. V. Poor, "Introducing the fast nonlinear Fourier transform," in *2013 IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, Vancouver, Canada, May 2013, pp. 5780–5784.
- [10] —, "Fast numerical nonlinear Fourier transforms," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6957–6974, Dec. 2015, also published at arXiv, Feb. 2014. [Online]. Available: <http://arxiv.org/abs/1402.1605>
- [11] —, "Fast inverse nonlinear Fourier transform for generating multi-solitons in optical fiber," in *IEEE Int. Symp. Info. Theory*, Hong Kong, China, Jun. 2015, pp. 1676–1680.

- [12] S. Wahls, S. T. Le, J. E. Prilepsky, H. V. Poor, and S. K. Turitsyn, "Digital backpropagation in the nonlinear Fourier domain," in *IEEE Int. Workshop on Signal Process. Advances in Wireless Commun. (SPAWC)*, Stockholm, Sweden, Jun. 2015, pp. 445–449, Also published at ArXiv on Apr. 2015, arXiv:1504.06598. [Online]. Available: <http://arxiv.org/abs/1504.06598>
- [13] G. Boffetta and A. R. Osborne, "Computation of the direct scattering transform for the nonlinear Schrödinger equation," *J. Comp. Phys.*, vol. 102, no. 2, pp. 252–264, Oct. 1992.
- [14] A. M. Bruckstein, B. C. Levy, and T. Kailath, "Differential methods in inverse scattering," *SIAM J. Appl. Math.*, vol. 45, no. 2, pp. 312–335, Apr. 1985.
- [15] J. Skaar and O. H. Waagaard, "Design and characterization of finite-length fiber gratings," *IEEE J. Quantum Electron.*, vol. 39, no. 10, pp. 1238–1245, Oct. 2003.
- [16] A. Rosenthal and M. Horowitz, "Inverse scattering algorithm for reconstructing strongly reflecting fiber Bragg gratings," *IEEE J. Quantum Electron.*, vol. 39, no. 8, pp. 1018–1026, Aug. 2003.
- [17] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: John Wiley & Sons, Inc., 1968.
- [18] A. Hasegawa and T. Nyu, "Eigenvalue communication," *IEEE J. Lightw. Technol.*, vol. 11, no. 3, pp. 395–399, Mar. 1993.
- [19] J. E. Prilepsky, S. A. Derevyanko, and S. K. Turitsyn, "Nonlinear spectral management: Linearization of the lossless fiber channel," *Opt. Exp.*, vol. 21, no. 20, pp. 24 344–24 367, Oct. 2013.
- [20] J. E. Prilepsky, S. A. Derevyanko, K. J. Bluw, I. Gabitov, and S. K. Turitsyn, "Nonlinear inverse synthesis and eigenvalue division multiplexing in optical fiber channels," *Phys. Rev. Lett.*, vol. 113, no. 1, pp. 013 901 (1–5), Jul. 2014.
- [21] S. T. Le, J. E. Prilepsky, and S. K. Turitsyn, "Nonlinear inverse synthesis for high spectral efficiency transmission in optical fibers," *Opt. Exp.*, vol. 22, no. 22, pp. 26 720–26 741, Nov. 2014.
- [22] E. Meron, M. Feder, and M. Shtaiif, "On the achievable communication rates of generalized soliton transmission systems," *arXiv:1207.0297*, pp. 1–13, Jul. 2012. [Online]. Available: <http://arxiv.org/abs/1207.0297>
- [23] S. T. Le, S. Wahls, D. Lavery, J. E. Prilepsky, and S. K. Turitsyn, "Reduced complexity nonlinear inverse synthesis for nonlinearity compensation in optical fiber links," in *The European Conf. Lasers Electro-Optics (CLEO)*, Munich, Germany, Jun. 2015.
- [24] H. Bülow, "Experimental assessment of nonlinear Fourier transformation based detection under fiber nonlinearity," in *European Conf. Opt. Commun.*, Sep. 2014, pp. 1–3.
- [25] H. Bülow, "Experimental demonstration of optical signal detection using nonlinear Fourier transform," *IEEE J. Lightw. Technol.*, vol. 33, no. 7, pp. 1433–1439, Apr. 2015.
- [26] H. Bülow, "Nonlinear Fourier transformation based coherent detection scheme for discrete spectrum," in *Opt. Fiber Commun. Conf. and Exposition*, Los Angeles, CA, United States, Mar. 2015, pp. 1–3.
- [27] Q. Zhang, T. H. Chan, and A. Grant, "Spatially periodic signals for fiber channels," in *2014 IEEE Int. Symp. Info. Theory (ISIT 2014)*, Jun. 2014, pp. 2804–2808.
- [28] Z. Dong, S. Hari, T. Gui, K. Zhong, M. I. Yousefi, C. Lu, P.-K. A. Wai, F. R. Kschischang, and A. P. T. Lau, "Nonlinear frequency division multiplexed transmissions based on NFT," *IEEE Photon. Technol. Lett.*, vol. 27, no. 15, pp. 1621–1623, Aug. 2015.
- [29] S. Civelli, L. Barletti, and M. Secondini, "Numerical methods for the inverse nonlinear Fourier transform," in *Tyrrhenian Int. Workshop on Digital Commun. (TIWDC)*, Florence, Italy, Sep. 2015, pp. 13–16.
- [30] V. Aref, H. Buelow, K. Schuh, and W. Idler, "Experimental demonstration of nonlinear frequency division multiplexed transmission," in *European Conf. Opt. Commun.*, Sep. 2015, pp. 1–3, Also published at ArXiv on Aug. 2015, arXiv:1508.02577. [Online]. Available: <http://arxiv.org/abs/1508.02577>
- [31] P. Kazakopoulos and A. L. Moustakas, "Nonlinear Schrödinger equation with random Gaussian input: Distribution of inverse scattering data and eigenvalues," *Phys. Rev. E*, vol. 78, no. 1, pp. 016 603 (1–7), Jul. 2008.
- [32] —, "NLSE soliton spectral efficiency for Gaussian input," Poster presentation at 2015 Munich Workshop on Information of Optical Fiber, Dec. 2015.
- [33] Q. Zhang and T. H. Chan, "A spectral domain noise model for optical fibre channels," in *2015 IEEE Int. Symp. Info. Theory (ISIT 2015)*, Jun. 2015, pp. 1660–1664.
- [34] A. Maruta, A. Toyota, Y. Matsuda, and Y. Ikeda, "Experimental demonstration of long haul transmission of eigenvalue modulated signals," in *Tyrrhenian Int. Workshop on Digital Commun. (TIWDC)*, Sep. 2015, pp. 28–30.